

PATENT ABSTRACTS OF JAPAN

(11)Publication number : 2001-147697

(43)Date of publication of application : 29.05.2001

(51)Int.Cl. G10L 15/10
G10L 15/00
G10L 19/00

(21)Application number : 11-329489

(71)Applicant : MATSUSHITA ELECTRIC IND CO
LTD

(22)Date of filing : 19.11.1999

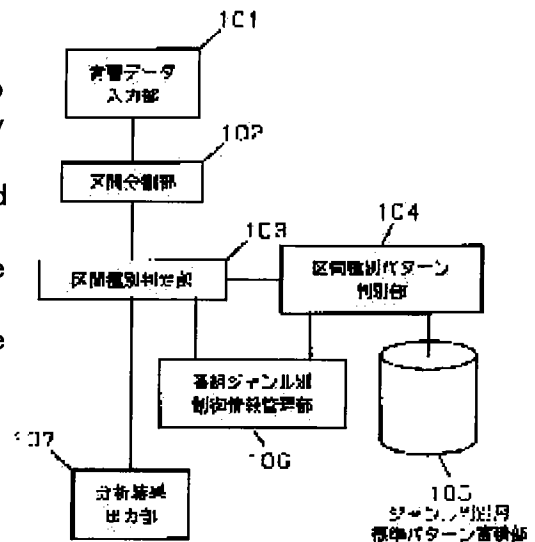
(72)Inventor : AKIMOTO TOSHIKI

(54) METHOD AND DEVICE FOR ACOUSTIC DATA ANALYSIS

(57)Abstract:

PROBLEM TO BE SOLVED: To detect a voice section with high precision even in the case of a mixed sound signal like a broadcasting program.

SOLUTION: The sound signal of a program is inputted to an acoustic data input part 101, and the feature quantity or the like is used to divide it into proper sections by a section division part 102. Acoustic data for every divided section is classified to three patterns of 'voice', 'non-voice', and 'mixture' by a discrimination part 103, and the genre of the program is determined by the pattern of classified acoustic data by a discrimination part 104. The start time, the end time, the degree of importance, etc., of the section are added on the basis of the determined genre to output the analysis result by an output part 107.



LEGAL STATUS

[Date of request for examination] 13.03.2003

[Date of sending the examiner's decision of rejection]

[Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]

[Date of final disposal for application]

[Patent number] 3757719

[Date of registration] 13.01.2006

[Number of appeal against examiner's decision of rejection]

[Date of requesting appeal against examiner's decision of rejection]

[Date of extinction of right]

(19) 日本国特許庁 (J P)

(12) 公 開 特 許 公 報 (A)

(11) 特許出願公開番号

特開2001-147697

(P2001-147697A)

(43) 公開日 平成13年5月29日 (2001.5.29)

(51) Int.Cl. ⁷	識別記号	F I	テームト* (参考)
G 1 0 L 15/10		G 1 0 L 3/00	5 3 1 N 5 D 0 1 5
15/00			5 5 1 G
19/00		9/00	J

審査請求 未請求 請求項の数12 O L (全 6 頁)

(21) 出願番号 特願平11-329489

(22) 出願日 平成11年11月19日 (1999. 11. 19)

(71) 出願人 000005821

松下電器産業株式会社

大阪府門真市大字門真1006番地

(72) 発明者 秋元 俊昭

大阪府門真市大字門真1006番地 松下電器
産業株式会社内

(74) 代理人 100097445

弁理士 岩橋 文雄 (外2名)

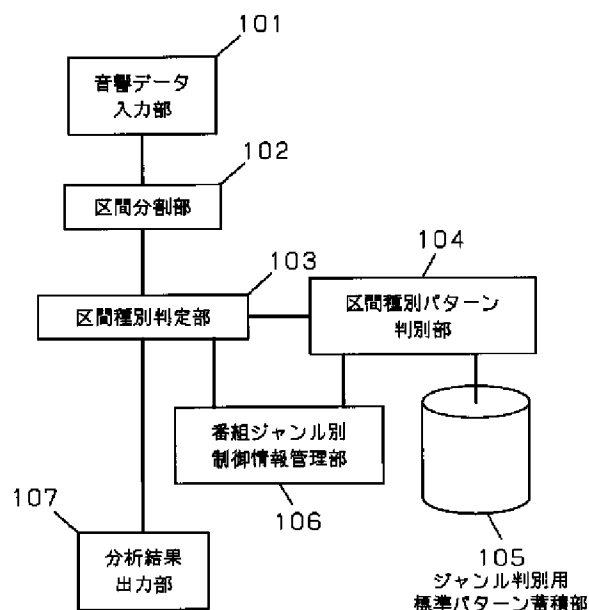
Fターム(参考) 5D015 AA06 CC05 DD01 FF06 HH23

(54) 【発明の名称】 音響データ分析方法及びその装置

(57) 【要約】

【課題】 放送番組などの混合された音声信号であっても、音声区間を精度良く検出すること。

【解決手段】 番組の音声信号が101に入力され(101)、特徴量などを使って、適当な区間に分ける(102)。分けた区間毎の音響データを「音声」「非音声」/「混合」の3つのパターンに分類し(103)、その分類された音響データのパターンにより、番組のジャンルを決定する(104)。決定されたジャンルに基づいて、区間のスタート時間、エンド時間、重要度、等を付加して分析結果として出力する(107)。



【特許請求の範囲】

【請求項1】 音響データを複数の区間に分割する音響データ分割手段と、前記区間の種別を判定する区間種別判定手段と、区間種別の統計処理結果を用いて番組のジャンルを判別する区間種別パターン判別手段と、番組のジャンルに応じた制御情報を管理する制御情報管理部とを備えたことを特徴とする音響データ分析装置。

【請求項2】 請求項1に記載の音響データ分析装置において、連続する同一種別区間を統合する区間統合手段を備えたことを特徴とする音響データ分析装置。

【請求項3】 音響データの自己相関係数を用いて音響データを複数の区間に分割することを特徴とする音響データ分割装置。

【請求項4】 請求項1に記載の音響データ分析装置において、音声を含む有音区間、音声を含まない有音区間、無音区間のいずれかの種別に判定する区間種別判定手段を備えたことを特徴とする音響データ分析装置。

【請求項5】 音響データを複数の区間に分割する音響データ分割手段と、前記区間の種別を判定する区間種別判定手段と、番組のジャンルに応じた制御情報を管理する制御情報管理部とを備えたことを特徴とする音響データ分析装置。

【請求項6】 音響データを複数の区間に分割する音響データ分割手段と、前記区間の種別を判定する区間種別判定手段とを備えたことを特徴とする区間種別パターン判別装置。

【請求項7】 音響データを複数の区間に分割する音響データ分割過程と、前記区間の種別を判定する区間種別判定過程と、区間種別の統計処理結果を用いて音響データのジャンルを判別する区間種別パターン判別過程と、ジャンルに応じた制御情報を使って区間種別を判定するジャンル適応型区間種別判定過程とを備えたことを特徴とする音響データ分析方法。

【請求項8】 請求項6に記載の音響データ分析方法において、連続する同一種別区間を統合する区間統合過程を備えたことを特徴とする音響データ分析方法

【請求項9】 音響データの自己相関係数を用いて音響データを複数の区間に分割することを特徴とする音響データ分割方法。

【請求項10】 前記区間種別判定過程において、音声を含む有音区間、音声を含まない有音区間、無音区間のいずれかの種別に判定することを特徴とする音響データ分析方法。

【請求項11】 音響データを複数の区間に分割する音響データ分割過程と、前記区間の種別を判定する区間種別判定過程と、ジャンルに応じた制御情報を管理する制御情報管理過程とを備えたことを特徴とする音響データ分析方法。

【請求項12】 音響データを複数の区間に分割する音響データ分割過程と、前記区間の種別を判定する区間種

別判定過程とを備えたことを特徴とする区間種別パターン判別方法。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】本発明はディジタル放送番組等のディジタルコンテンツ情報のスキミング、検索、フィルタリングを実現するのに必要な付加情報を自動的に抽出するための技術に相当する。

【0002】

【従来の技術】近年、ディジタル技術の発達によって、音や映像等のディジタルコンテンツ情報を大量に送信し、蓄積することができるようになった。大量のディジタルコンテンツ情報の中から、効率的に目的の情報を取り出すための方法として、スキミング、検索、フィルタリング等が知られているが、この方法を実現するためには、あらかじめディジタルコンテンツ情報の構造を解析し、構造に応じた付加情報を付ける必要がある。このような付加情報を人手に頼って付けていたのでは、コストがかかりすぎるため、自動的にディジタルコンテンツ情報の構造を解析し、構造に適した付加情報を付与する技術の研究開発が行われている。

【0003】例えば、USP-5,918,223(Method and article of manufacture for content-based analysis, storage, retrieval, and segmentation of audio information. June 29, 1999)では、音の分析方法と検索方法が特許登録されている。上記従来例は、音量、ピッチ、明るさ、帯域、メル周波数ケプストラム係数(MFCC)といった特徴量を統計処理し、音響情報の種別を判別するものである。これは、例えばインターネット上にある単一音源の音響ファイル(あるいはセグメント)を分類、検索するのに有効な手法である。また、上記技術を用いて、連続音をセグメントする方法を3種類示している。

【0004】1番目は、ユーザの提示する音の代表値(特徴ベクトルの統計量)との類似度がしきい値以上の領域とそれ以外とにセグメントする方法である。この方法によって、連続音は、2種類の領域に分割される。また、2番目は、ユーザの提示する音と全く同じ領域とそうでない領域にセグメントする方法である。最初の方法は、特徴ベクトルの統計量を代表値として類似度計算するのに対し、2番目の方法は、特徴ベクトルそのものを使って類似度を計算するところが大きな違いである。3番目の方法は、連続音の隣接する領域を比較し、類似度がしきい値以下となるポイント(シーンチェンジ)を検出するものである。

【0005】

【発明が解決しようとする課題】しかしながら、放送番組などの音響データを解析してみると、単一音源の音は少なく、音声と音楽が混合した音、観客の騒音とアナウンサーの声の混合した音等、複数の音が混在していることがわかる。このように、実際には混合音が多いため、

音の種別を判定するだけの手法は放送番組のようなコンテンツに対して実用的でないという問題点がある。

【0006】また、音響データは多様であるため、隣接する領域の類似度のみでシーンチェンジを検出すると、過剰に検出される。さらには、どのセグメントがその番組中重要であるかを判定することも困難である。

【0007】

【課題を解決するための手段】この課題を解決するために本発明は、音響データを意味のある区間に分割し、前記区間の重要度を判定する音響データ分析装置において、音響データを複数の区間に分割する音響データ分割手段と、前記区間の種別を判定する区間種別判定手段と、区間種別の統計処理結果を用いて音響データのジャンルを判別する区間種別パターン判別手段と、ジャンルに応じた制御情報を管理する制御情報管理部とを備えたものである。

【0008】本発明によれば、音響区間の単純な種別（音量の大きさ、音声が含まれる有音、音声が含まれない有音、無音）をいったん判別し、この音響区間の種別パターンによって、番組のジャンルを推定するので、ジャンルに適した混合音のモデルを設定できる。これにより、詳細な種別判定を行うことができるとともにやジャンルに適した重要度の判定を行うことができる。また、隣接する区間の特徴量の類似度でなく、隣接する区間の特徴量の変化パターン（増加／減少）、変化前の特徴量の値、変化後の特徴量の値、変化の大きさを利用することにより、過剰な区間検出を抑制することができる。

【0009】

【発明の実施の形態】本発明の請求項1に記載の発明は、音響データを複数の区間に分割する音響データ分割手段と、前記区間の種別を判定する区間種別判定手段と、区間種別の統計処理結果を用いて番組のジャンルを判別する区間種別パターン判別手段と、ジャンルに応じた制御情報を管理する制御情報管理部とを備えたことを特徴としており、番組のジャンルの推定、番組のジャンルに適した重要度判定を行うことができるという作用を有する。

【0010】本発明の請求項2に記載の発明は、請求項1に記載の音響データ分析装置において、連続する同一種別区間を統合する区間統合手段を備えたことを特徴としており、同一の分類区間が連続する場合、一つにまとめることができるという作用を有する。

【0011】本発明の請求項3に記載の発明は、音響データ分割手段において、音響データの自己相関係数を用いて音響データを複数の区間に分割することを特徴としており、音声区間の判別を精度良く、簡単に行うことができるという作用を有する。

【0012】本発明の請求項4に記載の発明は、請求項1に記載の音響データ分析装置において、音声を含む有音区間、音声を含まない有音区間、無音区間のいずれか

の種別に判定する区間種別判定手段を備えたことを特徴としており、ジャンルを判別する前の種別判定を高精度に実現でき、ジャンルの判別にも有効なパターンを生成するという作用を有する。

【0013】本発明の請求項5に記載の発明は、音響データを複数の区間に分割する音響データ分割手段と、前記区間の種別を判定する区間種別判定手段と、ジャンルに応じた制御情報を管理する制御情報管理部とを備えたことを特徴としており、ジャンル情報を用いて音響の分割区間の重要度を推定できるという作用を有する。

【0014】本発明の請求項6に記載の発明は、音響データを複数の区間に分割する音響データ分割手段と、前記区間の種別を判定する区間種別判定手段とを備えたことを特徴としており、番組のジャンルを推定することができるという作用を有する。

【0015】以下、本発明の実施の形態について、図を用いて説明する。

【0016】（実施の形態1）

ジャンル判定を行う場合のシステム

図1は、本発明の実施の形態1となる音響データ分析装置の構成図である。本実施の形態1の音響データ分析装置は、様々なフォーマットの音響データを番組単位に読み込む音響データ入力部101と、読み込んだ音響データを音響特徴量の変化情報を用いて分割する区間分割部102と、ジャンル不明の場合、音声区間、音声ありの混合音区間、音声なしの有音区間、無音区間のいずれかに区間の種別を判定し、ジャンルが判明した場合、ジャンルに適した種別に判定する区間種別判定部103と、ジャンルが既知の番組の区間種別判定結果を蓄積したジャンル判別用標準パターン蓄積部105と、区間種別パターンによって番組のジャンルを判別する区間種別パターン判別部104と、ジャンルに適した種別算出方法や重要度算出方法を管理する番組ジャンル別制御情報管理部106と、音響データの分析結果を出力する分析結果出力部107とから構成されている。

【0017】以上の音響データ分析装置の処理フローを図2に示す。まず音響データ入力部101を用いて音響データを入力し、区間分割部102を用いて入力音響データを複数の区間に分割する。以上が区間分割手段（step201）に相当する。次に、区間種別判定部103を用いて音声区間、音声ありの混合音区間、音声なしの有音区間、無音区間のいずれかに区間の種別を判定し、判定結果を区間種別パターン判定部104に出力する。以上が区間種別判定手段（step202）に相当する。区間種別パターン判定部104は、区間種別判定部103から出力される区間種別のパターンと、ジャンル判別用標準パターン蓄積部105に蓄積されているジャンルが既知の番組の区間種別パターンとを比較し、類似度の高い上位K個の番組を選択する。選択した番組のジャンルで最も多いジャンルを音響データ入力部101か

ら入力した音響データのジャンルとする。このジャンル推定結果を番組別制御情報管理部106に入力する。以上が区間種別パターン判別手段（step203）に相当する。番組ジャンル別制御情報管理部106は、番組ジャンルに適した区間種別情報と重要区間判定情報を制御情報として区間種別判定部103に出力する。区間種別判定部103は、ジャンル別制御情報を用いて、区間種別や区間重要度を計算し、分析結果出力部107へ出力する。以上がジャンル適応型区間種別判定手段（step204）に相当する。分析結果出力部107は分析結果を適当なフォーマットに変換して出力する。例えば、区間のスタート時間、区間のエンド時間、区間重要度の値、区間種別番号を行単位に出力する。

【0018】具体的な区間分割手段（step201）を図3のフローチャート図を用いて説明する。音響データを入力として自己相関分析を行い、正規化した1次の自己相関係数を出力する（step301）。正規化した1次の自己相関係数を用いて音響データの分割位置を検出し、分割位置のランク付けをする（step310）。分割位置のランクに応じて音響データを分割する（step320）。さらに分割位置検出手段（step310）は、音響データの全ての位置で前後区間の正規化1次自己相関係数の最大値を計算し（step311）、隣接する分割区間の最大値の差を計算するとともに（step312）、時間方向に対して最大値が増加するパターン（立ち上がりの位置）を見つける（step313）。さらに、立ち上がり位置の変化の大きさに応じてN段階のランク付けを行い（step314）、一定ランク以上の分割位置の前後数秒以内（例えば2秒以内）に一定ランク以下の分割位置があるかどうかの判定を行い（step315）、一定ランク以下の分割位置が見つければ、その分割位置のランクを1ランク下げる（step316）。

【0019】具体的な区間種別判定判別手段（step202）を図4のフローチャート図を用いて説明する。区間単位にしきい値以下の1次自己相関係数の割合、1次自己相関係数の平均、音量のピーク等の音響特徴量を計算し（step401）、しきい値以下の1次自己相関係数の割合が α 以上 β 以下であるかの判定を行い（step402）、条件が合えば音声区間とみなす。また、しきい値以下の1次自己相関係数の割合が α 以下であるかの判定を行い（step403）、条件が合えば音声ありの混合音区間とみなす。さらに、上記ステップで条件の合わなかった区間の内、音量ピークが γ 以上かどうかの判定を行い（step404）、条件が合えば音声なしの有音区間とみなす。以上のステップでいずれの条件にも合わなかった区間を無音区間とみなす。

【0020】具体的な区間種別パターン判別手段（step203）を図5のフローチャート図を用いて説明する。番組単位に、各区間種別の割合、区間長が10秒以

上／1分以上／3分以上の区間の割合、区間長が10秒以下／1分以下／3分以下の区間の割合、さらには区間種別毎の区間長の割合等の番組種別統計量を計算し（step501）、前ステップで求めた番組種別統計量を使って、ジャンル判別用標準パターンとの類似度を計算し、類似度の高い上位N個を選択する（step502）。選択したN個の番組中、最も多いジャンルを当該音響データのジャンルとする。ジャンル数が同じ場合は、最も類似度の高い番組のジャンルとする（step503）。

【0021】具体的なジャンル適応型区間種別判定手段（step204）を図6のフローチャート図を用いて説明する。あらかじめ用意したジャンルに適した制御情報（区間種別情報や重要区間情報）を読み込み（step601）、ジャンルに適した特徴量を計算し（step602）、区間種別の判定及び区間種別に応じた重要度を計算する（step603）。ジャンルに適した区間種別情報とは、例えば、音楽番組では、区間種別として音楽区間、トーク区間があり、音楽区間の特徴を判定するための特徴量とそのしきい値、トーク区間と判定するための特徴量とそのしきい値に相当する。区間種別を判定するための特徴量として、音量、ピッチ、明るさ、帯域、メル周波数ケプストラムの他に、自己相関係数を用いる。例えば、音楽番組というジャンルの音楽区間を判定するためには、分割区間単位に正規化1次自己相関係数が0.01以下となる部分区間の割合が0.1以下かつ正規化1次自己相関係数が0.1以上という条件を設定する。重要区間判定情報とは、例えば、音楽番組では、音楽区間の重要度が高く、トーク区間の重要度が低いというような情報に相当する。

【0022】（実施の形態2）

ジャンル情報が入力される場合

図7は、音響データとともに、ジャンル情報が入力される場合の実施の形態となる音響データ分析装置のシステム構成図である。本実施の形態1の音響データ分析装置は、様々なフォーマットの音響データを番組単位に読み込む音響データ入力部701と、読み込んだ音響データを音響特徴量の変化情報を用いて分割する区間分割部702と、ジャンルに適した種別に判定する区間種別判定部703と、ジャンルに適した種別算出方法や重要度算出方法を管理する番組ジャンル別制御情報管理部704と、音響データの分析結果を出力する分析結果出力部705とから構成されている。

【0023】以上の音響データ分析装置を構築するためには、まず音響データ入力部701を用いて音響データを入力し、区間分割部702を用いて入力音響データを複数区間に分割する。次に、区間種別判定部703はジャンルをこのジャンル推定結果を番組別制御情報管理部704に出力する。番組別制御情報管理部704は、ジャンルに適した区間種別計算方法や重要区間判定方法を

制御情報として区間種別判定部 703 に出力する。区間種別判定部 703 は、ジャンル別制御情報を用いて、区間種別や区間重要度を計算し、分析結果出力部 705 へ出力する。分析結果出力部 705 は分析結果を適当なフォーマットに変換して出力する。

【0024】実施の形態 1 との違いは、ジャンル情報があらかじめわかっているかどうかである。実施の形態 2 では、ジャンル情報があらかじめわかっている場合を想定しており、実施の形態 1 において、ジャンルを推定する装置を除いた構成となる。

【0025】

【発明の効果】以上のように本発明の音響データ分析装置及び方法は、いったん音響区間の単純な種別（音量の大きさ、音声が含まれる有音、音声が含まれない有音、無音）を判定し、この判定結果によって番組のジャンルを推定しているので、分割区間の混合音の種別判定や重要度の判定を行うことができるという効果がある。

【図面の簡単な説明】

【図 1】本発明の実施の形態 1 となる音響データ分析装

置の構成図

【図 2】音響データ分析装置の処理全体を表すフローチャート

【図 3】区間分割手段のフローチャート

【図 4】区間種別判定手段（1）のフローチャート

【図 5】区間種別パターン判別手段のフローチャート

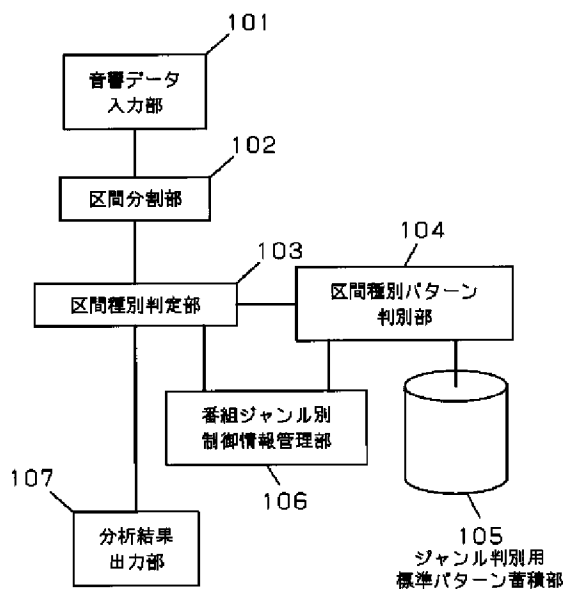
【図 6】ジャンル適応型区間種別判定手段のフローチャート

【図 7】本発明の実施の形態 2 となる音響データ分析装置のシステム構成図

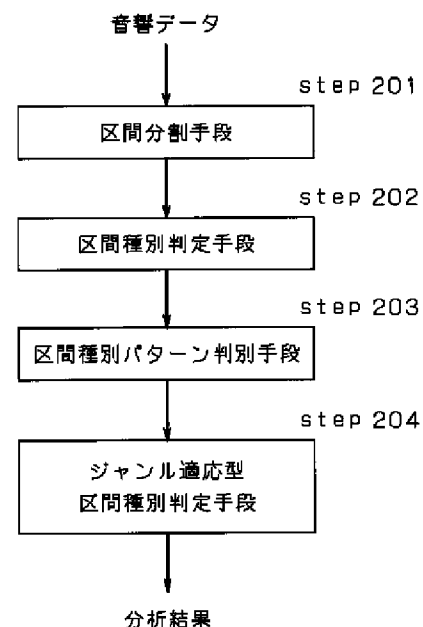
【符号の説明】

- 101 音響データ入力部
- 102 区間分割部
- 103 区間種別判定部
- 104 区間種別パターン判別部
- 105 ジャンル判別用標準パターン蓄積部
- 106 番組ジャンル別制御情報管理部
- 107 分析結果出力部

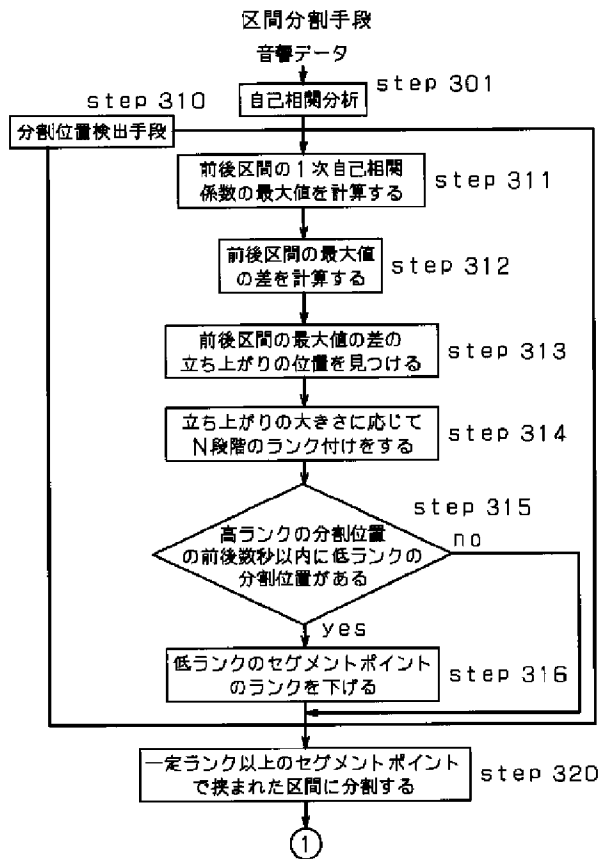
【図 1】



【図 2】

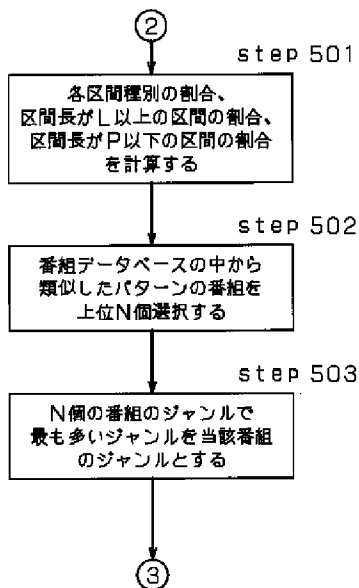


【図3】



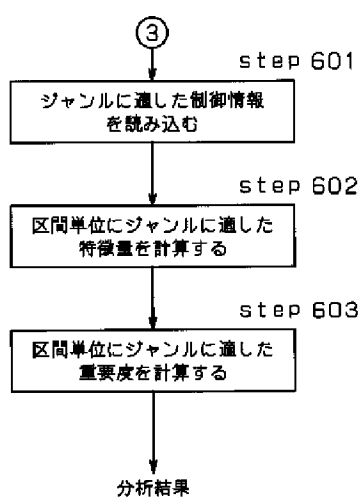
【図5】

区間種別パターン判別手段

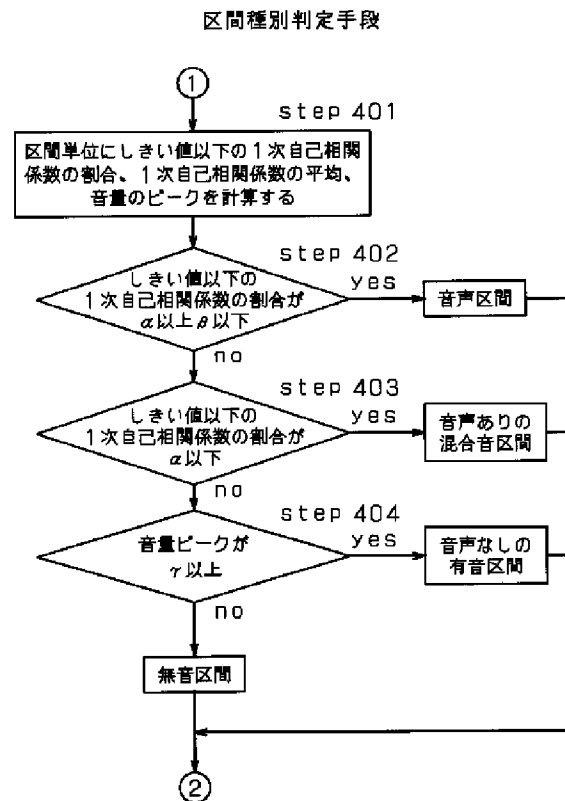


【図6】

ジャンル適応型区間種別判定手段



【図4】



【図7】

